

電子計算機によるX線回折データの検索

幾志新吉

国立防災科学技術センター第3研究部情報処理研究室

Information Retrieval of X-Ray Diffraction Data with Digital Computer

By

S. Kishi

National Research Center for Disaster Prevention, Tokyo

Abstract

The present report describes a method of identification of crystalline materials, using a digital computer. The pattern of the optical diffraction intensities for the interplanar spacings of various crystalline structures belonging to a crystalline material can be obtained by radiating X-ray upon the crystalline material. This pattern is the characteristic of the substance, its own fingerprint, so to speak. Accordingly, if we have a library of such fingerprints, we can identify an unknown substance by finding the same or similar pattern in the library.

In our information retrieval, an A. S. T. M. Diffraction Data File was used as the library, from which we had recorded about 13,000 substances on magnetic tape, and the operation of finding is conducted with the computer TOSBAC-3400.

When the substance sought is a single substance, the process is simple, because we have only to find a single card. However, when it is a mixture, the process is quite different. By the first retrieval, the possible substances are reduced to about 1%. Next, for the second retrieval, a hypothetical number of constituents is set beforehand, by applying the method of least squares to the diffraction intensities, in order to find the combinations of the constituent substances of the mixture.

This experiment deals with the substances excepting the extremely complicated minerals and those, such as clay, having broad or blunt peaks. Therefore, the further experiments are expected to be proceeded, from the practical point of view, by limiting the object substances to minerals, for instance, aiming at developing some information retrieval methods proper to the object substances.

1. 概 要

結晶性物質にX線を照射すると、その結晶性物質に応じたいろいろな結晶構造の格子面間隔に対する光学的回折強度のパターンが得られる。このパターンは物質固有のものであり、いわば物質の指紋である。したがって多数の物質の指紋台帳があれば、未知の物質の指紋を取って、台帳の中から同じものまたは似たものを捜すことによって、未知の物質が何であるかを決定することができる。

われわれの情報検索では、指紋台帳としてA.S.T.M. (American Society for Testing and Materials) カードファイルを用い、そこから物質約 13,000種類 のデータを計算機の磁気テープに収め、捜す操作を電子計算機 TOSBAC-3400で行なった。

被検索物質が単体の場合には、磁気テープには入っているカードの1枚を捜し当てればよいのだから簡単であるが、混合物となると根本的に方法を変えねばならず、まず一次検索で候補となる物質を1%ほどにしぼり、次に二次検索として、あらかじめ成分数を仮定したうえで、それらの回折強度に対して最小二乗法を適用し、どの物質とどの物質が混じっているかを見いだすようにした。

2. データ

2.1 A.S.T.M.カードの実例を図-8に示す。

右半分がわれわれの用いたデータで、それらは結晶構造の格子面間隔 d (単位 λ) と、最強の回折強度 I_0 に対するその格子面間隔に対応する回折線強度 I が百分比 I/I_0 で表わされ、必ず100を一つまたはそれ以上含んだものである。 I/I_0 欄には後述のチャートにおけるブロードを示す記号の付いていたものもあったが、われわれは d と I/I_0 だけで物質の識別が可能であるもののみを検索の初段階としてねらったので、これは無視した。また TOSBAC-3400 は1語 24 bit であるから、対応する d と I/I_0 の数値を1語に詰め込むと過不足なく好都合であった。1物質について平均約30対の d と I/I_0 のかきこみが終わったら、次にカード左上の物質番号を1語として入れた。

以上のデータは数名で半年かかって作製したが、その数は物質数で13,067種、語数で344,261語であった。

ここで、次に述べるチャートとの比較のために、A.S.T.M.カードの数値パターンをモデル化してみ

ると図-1のようになる。

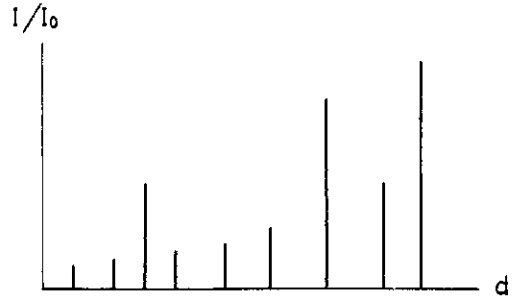


図-1 A.S.T.M.カードのパターン

2.2 結晶性の未知物質をX線回折装置にかけると、たとえば図-9、図-10、図-11のようなチャートが得られる。横軸は入射X線ビームと結晶の格子面とのなす角 θ の2倍角である。よく知られているように、Braggの反射条件

$$2d \sin \theta = n\lambda$$

または

$$d = n\lambda / 2 \sin \theta$$

が満足されるときに回折線のピークが現われる。ここに n は結晶の性質によってきまる整数値で、 λ は入射X線の波長であって、X線を発生させるターゲットの元素に固有なものである。

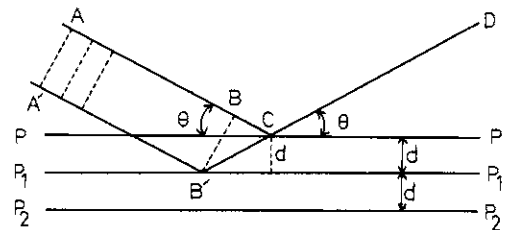


図-2 Bragg反射の説明図

次に図-2によって簡単にこれを示す。波長 λ の入射X線 $AB, A'B'$ が格子面 P, P_1 で CD の方向に反射されるものとする。路程差 $B'C - BC$ が波長の整数倍に等しくなるときに、反射X線は強めあう。

すなわち

$$B'C - BC = n\lambda,$$

また簡単な三角法から、

$$B'C = d / \sin \theta$$

および

$$BC = BC \cos 2\theta$$

$$= d (\cos 2\theta) / \sin \theta.$$

あとの2式を第1式に入れて

$$2d \sin \theta = n\lambda$$

を得る。

したがって、チャートからピークの現われる角度を読み取って Bragg の式から d を求め、一方最強強度に対する各々の強度比を読み取れば、未知物質に関する図-1のパターンが得られる。

しかし、実際にはチャートにおけるピークの位置は測定条件の差異によって変動するので、ピークの位置は適当な幅(区間 D)をもたせて読み取ることとした。この幅については混合物のところでも一言する。

そしてチャートにおける山型を図-3のように長方形で近似させることにする。

Braggの式によって、図-3のモデルの横軸の 2θ を d に変換したのが図-4である。

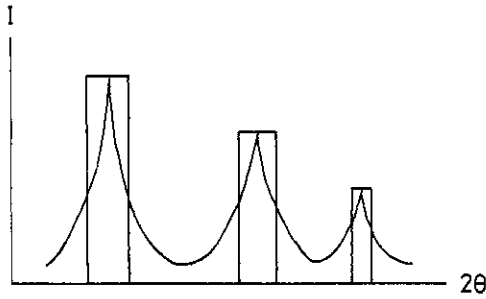


図-3 ピークの長方形近似

以下の検索は、図-1のモデルと図-4のモデルの相似性を考えればよいわけである。

もちろんピークの高さ I/I_0 の方もピークの位置 d 以上に誤差を含むが、この処理は以下の計算の過程に含まれている。

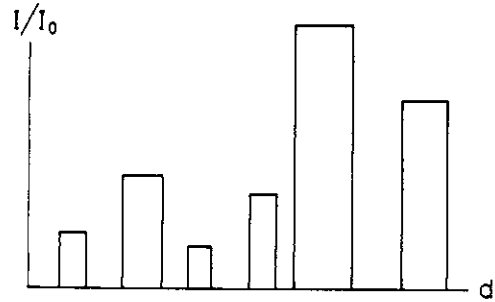


図-4 被検索物質のパターン

3. 単体の検索

いろいろな方法が考えられるが、検索を2段階にわけるといふ点で混合物の検索と同じ考え方でできる次の方法をとった。

なお、区間といった場合、それを 2θ 軸における区間(図-3)と考へても、 d 軸における区間(図-4)と考へても、以下の叙述において本質的な相違はないから、この2種類の区間についてあえて別の記号は使わなかった。

3.1 一次検索として、未知物質のチャートからピークの高さを考慮して読み取った区間の集合 $\{D_k\}$ (ここに $k=1, 2, \dots, m$; m は読み取り区間の数)の外に、ある一定値 I_p 以上の強度 (I/I_0) をもつ物質を、A.S.T.M.カードの中からふるい落としてゆく。なぜならばA.S.T.M.カードの中の物質がその区間外にも大きなピークをもつ物質であれば、それがチャートに現われなければならないからである。

そこで、測定条件の差異を考慮しても、現われなければならないと思われる最低の強度を I_p に選べばよいであろう。

理論的には、 I_p を小さく選ぶほど、ふるい落としの効果は上がるが、あまり小さくすると、何かの原因でA.S.T.M.カードに記入された標準としている回折像に比べて、今問題としている未知物質のそのときの回折像がかなり小さくて、ノイズの中に埋もれている場合、肝心の物質のカードをふるい落としてしまう危険がある。

いくつか検索してみたところでは、単体の場合、

I_p は40ぐらいにしておけばまず安全である。

この一次検索については混合物のところで再び述べる。

3.2 次に、二次検索として、ふるい落とされずに残ったもの n 個について、回折強度に関する検索をする。区間 D_k において、未知物質の強度 $(I/I_0) x_k$ と、A.S.T.M.カードの物質 j の強度 P_{jk} が対応している。

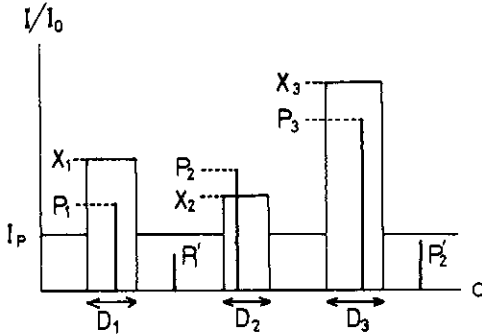


図-5 単体の検索

P_{jk} と x_k の差の絶対値を全区間について加える。

$$\epsilon_j = \sum_{k=1}^m |x_k - P_{jk}|$$

($j = 1, \dots, n$)

この ϵ_j の小さな物質ほど、未知物質と似ているといふことができる。

また、 $\{x_k\}$ と $\{P_{jk}\}$ の相関係数 ρ_j を用いて検索してみたが、いずれにしても後述するよりよい結果が得られた。

所要時間は約1分半でそのほとんどが磁気テープの回転時間である。

4. 混合物の検索

4.1 混合物の成分が3物質以上の場合でも全く同様であるから、しばらく成分数が二つの場合について述べる。

物質 X, Y が混合しているとき、単体における回折強度 I/I_0 をそれぞれ x, y 、混合物のそれを z とすれば、 α, β を1より小さい正の定数として、

$$z = \alpha x + \beta y$$

$$(0 \leq x, y, z \leq 100)$$

なる関係がある。

以後この α, β を便宜上混合比と呼ぶことにす

る。いま、混合物すなわち未知物質を構成する物質を X, Y と仮定した場合、最小二乗法によって、

$$\epsilon_{xy} = \sum_{k=1}^m (ax_k + by_k - z_k)^2$$

を最小にする a, b を決めることができ、同時に ϵ_{xy} を計算することができる。

4.2 そこで、前述のような一次検索に合格した候補物質 n 個の、すべての2組 ${}_n C_2 = n(n-1)/2$ 個について ϵ_{xy} を計算し、その小さい組合せの物質ほど、未知物質の成分としての可能性があると考えられる。

各区間にはいった成分のピークは、実際には図-6のようにずれているであろうが、このズレは無視して一致しているものとみなし、混合物のピークを、それらの線型結合で近似するのである。換言すれば、成分のピークの強度が各区間において一様に分布しているものと仮定する。つまり図-1の棒についても、図-4のように長方形にしてしまうのである。

図-7はこの事情を説明したものである。

4.3 ところで、混合物の一次検索では、単体の場合に比べて、山の幅 D' およびふるい落としの基準値 I_p はともに大きめにとるべきである。 D' については、山の頂点のズレだけでなく、図-6からわかるように、混合物の山の中腹に成分のピークが隠れている可能性も考慮しなければならない。

全区間 $\{D_k\}$ がきまったら、 $\{D_k\}$ 外に I_p 以上のピークをもつ物質をふるい落とすわけであるが、この I_p をきめる条件は二つあり、その一つは次のような考察に基づくものである。

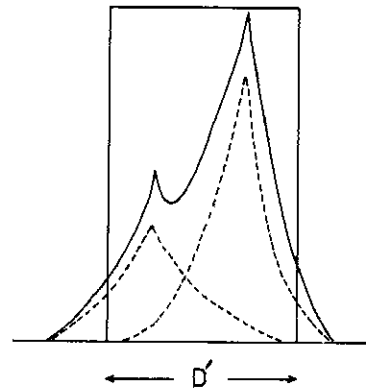


図-6 混合物と成分のピーク

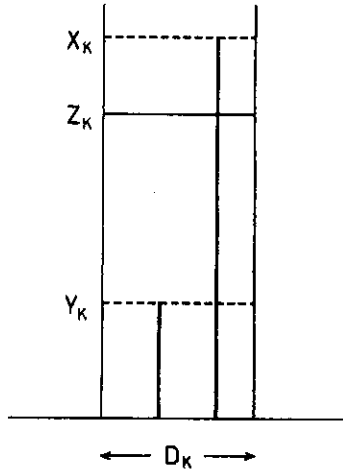


図-7 $z_k = ax_k + by_k$

単体における $I (\leq 100)$ という強度は、それが、 α_0 (前述の意味での混合比で $0 < \alpha < 1$) の割合で混じっている混合物においては、 $I\alpha_0$ となるから、単体における強度 I のピークが、混合物のチャートの $\{D_k\}$ 外において検出されるためには、チャートにおけるノイズレベルの強度を N とすれば、

$$I\alpha_0 > N$$

でなければならない。

ここでノイズレベルというのは、小さなピークかノイズかの区別がはっきりしないような強度レベルのことであって、多分に人為的に定めるものである。

$\{D_k\}$ の外に、 $I > N/\alpha_0$ なるピークをもつ単体ならば、それがチャートに現われないはずはない。ところが実際にはそのピークはチャートに現われていないのであるから、このような単体は成分ではあり得ない。

そこで混合比 α_0 の単体に関しては、

$$I_p = N/\alpha_0$$

とすればよいであろう。

I_p をきめるときのもう一つの条件は、後述するように、単体の場合と違って、一次検索による合格物質数を制限しなければならないということである。

このようなかねあいから I_p をきめるのであるが、この一次検索が有効に行なわれるか否かがわれわれの方法全体のかぎであるということができよう。

一次検索の結果残った物質 n 個の、各区間に対応する強度の配列

$$\{x_{j1}, x_{j2}, \dots, x_{jm}\} \\ (j = 1, \dots, n)$$

は以下の計算のために、記憶装置に記憶させておく。

一次検索の所要時間は、単体の検索のときと同じく約1分半である。

4.4 ここで、二次検索の計算の結果定まった、ある物質の混合比を α とし、仮定混合比 α_0 との関連を考えてみる。

$\{D_k\}$ の外にある、A.S.T.M.カードのピークの最大値を x_0 とする。

i) $\alpha < \alpha_0, x_0 > I_p$ のとき、 $\alpha x_0 < N$ となりうる。すなわち α が小さい場合には一次検索で落とされたものの中にも成分となりうるものがあるかもしれないということである。

したがって、この方法で混合比の小さな成分まで検索することは無意味である。

ii) $\alpha > \alpha_0, x_0 < I_p$ のとき、 $\alpha x_0 > N$ となりうる。

このような成分は最終的におとす方法も考えられる。

iii) $\alpha > \alpha_0, x_0 > I_p$ のときは常に、 $\alpha x_0 > N$ だから、一次検索であるいおとしたことに問題はない。

iv) $\alpha < \alpha_0, x_0 < I_p$ のときも常に、 $\alpha x_0 < N$ だから、問題はない。

4.5 さて、二次検索における最小二乗法を一般化してみよう。成分数 r 、区間数 m 、 j 成分の強度を

$$x_{j1}, x_{j2}, \dots, x_{jm}, \\ (j = 1, \dots, r)$$

混合物の強度を

$$y_1, y_2, \dots, y_m$$

とする。 m 個の近似式

$$y_k = a_1 x_{1k} + \dots + a_r x_{rk} \\ (k = 1, \dots, m)$$

において、

誤差の平方和

$$\sum_{k=1}^m [(a_1 x_{1k} + \dots + a_r x_{rk}) - y_k]^2$$

を最小にする a_j を決めるという条件から、正規方程式 (r 元連立一次方程式)

$$\begin{pmatrix} \Sigma x_{1k}^2 & \dots & \Sigma x_{1k} x_{rk} \\ \vdots & \ddots & \vdots \\ \Sigma x_{rk} x_{1k} & \dots & \Sigma x_{rk}^2 \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} = \begin{pmatrix} \Sigma x_{1k} y_k \\ \vdots \\ \Sigma x_{rk} y_k \end{pmatrix}$$

が得られる。ここに Σ は $\sum_{k=1}^m$ の意である。この係数行列は対称行列であって、その要素は r 個の物質の任意の2組についての強度の積和である。

候補物質 n 個から、仮定成分 r 個をとるすべての組合せ ${}_n C_r = n! / (n-r)! r!$ に対して、上の連立一次方程式を解くわけであるが、そのつど積和を計算するのはむだである。そこで、前もって、 n 個から重複を認めて2個とる組合せ

${}_n H_2 = {}_{n+1} C_2 = (n+1)n/2$ 個の積和を計算し、

$$\begin{matrix} \Sigma x_{1k}^2 & \Sigma x_{1k} x_{2k} & \dots & \Sigma x_{1k} x_{nk} \\ \Sigma x_{2k}^2 & \dots & \dots & \Sigma x_{2k} x_{nk} \\ \dots & \dots & \dots & \dots \\ \Sigma x_{rk}^2 & \dots & \dots & \dots \end{matrix}$$

のように順序よく記憶装置に格納しておけば、ある正規方程式の係数行列は、対応する積和を取り出すだけでえられる。

1個の積和を1語に格納すると、 $n=100$ のとき5050語を要するから、全体の語数との関係でこれを限度とした。 ${}_{100} H_2$ の積和計算の所要時間は区間数 m が30ぐらいのとき約20秒である。

4.6 ところで、 n 個から r 個とる組合せの数というのは、実際にどのぐらいの値か計算してみよう。 n は50~100、 r は2~5ぐらいとするから、

$$\begin{aligned} {}_{50} C_2 &= 1225, \\ {}_{50} C_3 &= 19600, \\ {}_{50} C_4 &= 230300, \\ {}_{50} C_5 &= 2120000, \\ {}_{100} C_2 &= 4950, \\ {}_{100} C_3 &= 161700, \\ {}_{100} C_4 &= 3900000, \\ {}_{100} C_5 &= 75300000. \end{aligned}$$

これからわかるように、 ${}_n C_r$ は n が大きくふえてもその割りにふえないが、 r が一つふえただけで急激に増大する。 n, r がこのぐらいの範囲ならば一般に、

$$\begin{aligned} {}_n C_{r+1} &\approx \frac{n}{r+1} {}_n C_r, \\ {}_{n+m} C_r &\approx \left(\frac{n+m}{n}\right)^r {}_n C_r \end{aligned}$$

がなりたつ。

所要時間は ${}_{100} C_2$ で約20秒、 ${}_{100} C_3$ で15分であった。この割合でゆくと ${}_{100} C_4$ では数時間かかることになる。仮に ${}_{100} C_2$ が1秒だとしても ${}_{100} C_4$ で10分、 ${}_{100} C_5$ では4時間もかかる。計算機の演算速度には限度があり、このように急上昇する組合せについての正規方程式を解くことは時間的に不可能である。また、成分が多いと平均的に混合比は小さくなり、先述のごとく混合比の小さなものは意味がないから、成分数はあまり多くはできない。

このようなわけで、われわれは $r=5$ までのプログラムを作ることにした。

4.7 混合物の成分のうちいくつかは既に判明していて、残りを捜したい場合がある。また $r=4, 5$ ぐらいの混合物でも、その主成分は $r=2, 3$ のプログラムで大体判明するから、さらに残りの成分を r が大きいときのプログラムで検索したい場合もある。このようなときのために、 r 個を検索するプログラムでは、物質番号を指定することによって、1, ..., $r-1$ 個を固定できるようにになっている。

5. 検索例

5.1 単体

図-9は calcite のチャートで、対応する A.S. T.M.カードの番号は、5-0586 (図-8) であるが、これを未知の単体物質として検索してみた結果を、誤差

$$\Sigma |x_k - p_{jk}|$$

の小さな順に上位5個を列挙すると表-1のようになる。

図-10は andradite 10-0288 である。結果を表-2に示す。

いずれも山の中腹を区間の両端として読み取り、 $I_p = 40$ とした。読み取りピークの数はいずれも14および18である。またターゲットは Cu で $\lambda = 1.5405 \text{ \AA}$ である。

5-0586

d	3.04	2.29	2.10	3.86	CaCO ₃				
I/I ₀	100	18	18	12	CALCIUM CARBONATE CALCITE				
Rad. Cu	λ 1.5405	Filter		d Å	I/I ₀	hkl	d Å	I/I ₀	hkl
Dia.	Cut off	Coll.		3.86	12	102	1.297	2	218
I/I ₀		d corr. eba.†		3.035	100	104	1.284	1	306
Ref. SWANSON AND FUYAT, NES CIRCULAR 539, VOL. II, 51 (1953)				2.845	3	006	1.247	1	220
Sys. HEXAGONAL		S.G. D _{6h} - R3c		2.495	14	110	1.235	2	1.1.12
a ₀ 4.989	b ₀	c ₀ 17.062	A	2.285	18	113	1.1795	3	2.1.10
α	β	γ	Z 6	2.095	18	202	1.1538	3	314
Ref. IBID.				1.927	5	204	1.1425	1	226
θ ₀ 1.487	n ₀ 1.659	ε ₀	Sign -	1.913	17	108	1.1244	<1	2.1.11
ZV	D _x 2.71imp	Color		1.875	17	116	1.0613	1	2.0.14
Ref. IBID.				1.626	4	211	1.0473	3	404
SAMPLE FROM MALLINCKRODT CHEM. WORKS. SPECT. ANAL. : <0.1% Sr; <0.01% Ba; <0.001% Al, B, Cs, Cu, K, Mg, Na, Si, Sn; <0.000% Ag, Cr, Fe, Li, Mn. X-RAY PATTERN AT 260C				1.604	8	212	1.0447	4	138
REPLACES 1-0837, 2-0623, 2-0629, 3-0569, 3-0593, 3-0596, 3-0612, 4-0636, 4-0637				1.587	2	1.0.10	1.0352	2	0.1.16
				1.525	5	214	1.0352	2	1.1.15
				1.518	4	208	1.0234	<1	1.2.13
				1.510	3	119	1.0118	2	3.0.12
				1.473	2	215	0.9895	<1	231
				1.440	5	300	.9646	1	322
				1.422	3	0.0.12	.9782	1	1.0.17
				1.356	1	217	.9767	3	2.1.14
				1.339	2	2.0.10	.9655	2	234

図-8 calciteのA.S.T.M.カード

表-1 calciteの検索結果

カード番号	誤差	化学式	物質名または鉱物名
05-0586	74	CaCO ₃	calcite
11-0619	177	CaNb ₂ O ₆	fermite
12-0430	185	Eu ₆ Sm ₂ PO ₄	europium samarium phosphate
11-0026	185	β-Zn ₂ P ₂ O ₇	beta zinc pyrophosphate
01-0829	190	Ca(CNS) ₂ ·3H ₂ O	calcium thiocyanate trihydrate

表-2 andraditeの検索結果

カード番号	誤差	化学式	物質名または鉱物名
10-0288	98	Ca ₃ Fe ₂ Si ₃ O ₁₂	andradite
03-1135	188	3CaO·Fe ₂ O ₃ ·3SiO ₂	andradite (topazolite)
13-0014	282	Cd ₃ Al ₂ Ge ₃ O ₁₂	cadmium aluminum germanate
10-0265	298	Ca ₃ Al ₂ (GeO ₄) ₃	calcium aluminum germanate
03-1134	322	3CaO·Fe ₂ O ₃ ·3SiO ₂	andradite

なお、図-9のチャートを読み取って図-4の
パターンに変換した値は、

区間 D	I/I ₀
3.897 ~ 3.798 Å	7
3.055 ~ 2.996	100
2.506 ~ 2.466	9
2.291 ~ 2.257	17
2.102 ~ 2.074	16
1.916 ~ 1.894	19
1.879 ~ 1.857	17
1.627 ~ 1.612	3
1.606 ~ 1.589	5
1.527 ~ 1.500	4
1.445 ~ 1.433	4

これと図-8のカードの値が似ていたわけである。

5.2 混合物

図-11はα-Al₂O₃とZnOの混合物の例である。
読み取り値は次のとおりである。

区間 D	I/I ₀
25.4° ~ 25.8°	10
31.6 ~ 32.0	56
34.4 ~ 34.8	39
35.1 ~ 35.4	20
36.2 ~ 36.4	100
37.7 ~ 38.0	9
43.3 ~ 43.7	21
47.5 ~ 47.9	23
52.5 ~ 52.9	8
56.5 ~ 56.9	35

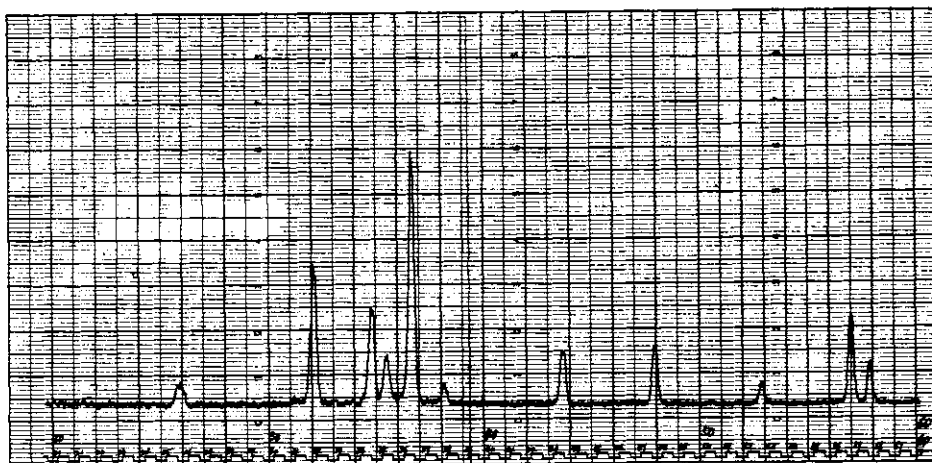


図-11 ZnO + Al₂O₃ のチャート

$I_p = 40$ とすると一次検索で82個残る。チャートから、 $N = 4$ として混合比は $\alpha_0 = N/I_p = 0.1$ ままで有意義である。

二次検索の結果は次のようになった。

誤差 $\Sigma (ax_k + by_k - z_k)^2$ の小さい順に上位5個を示すと、成分数 $r = 2$ としたとき表-3となる。

表-3 ZnO + Al₂O₃ の検索結果 ($r = 2$)

成分番号1	成分番号2	混合比1	混合比2	誤差
5-0664 (ZnO)	5-0712 (α -Al ₂ O ₃)	0.884	0.199	325.6
5-0664 (ZnO)	10-0173 (α -Al ₂ O ₃)	0.884	0.199	330.7
2-1227 (α -Al ₂ O ₃)	5-0664 (ZnO)	0.179	0.884	383.9
5-0664 (ZnO)	6-0662 (Os)	0.884	0.236	1032.0
1-1244 (Zn)	5-0664 (ZnO)	0.234	0.850	1100.0

表-4 ZnO + Al₂O₃ の検索結果 ($r = 3, \text{ZnO}$ 固定)

成分番号1	成分番号2	成分番号3	混合比1	混合比2	混合比3	誤差
5-664 (ZnO)	6-534 (NaPt ₃ O ₄)	5-712 (α -Al ₂ O ₃)	0.884	0.023	0.193	321.7
5-664 (ZnO)	6-662 (Os)	5-712 (α -Al ₂ O ₃)	0.884	0.023	0.190	322.2
5-664 (ZnO)	3-195 (MgFeSiO ₄)	5-712 (α -Al ₂ O ₃)	0.884	0.021	0.193	323.0
5-664 (ZnO)	10-117 (Li ₃ ZrF ₇)	5-712 (α -Al ₂ O ₃)	0.884	0.016	0.194	323.8
5-664 (ZnO)	8-249 (Li ₂ TiO ₃)	5-712 (α -Al ₂ O ₃)	0.884	0.016	0.194	323.8

5-664 は ZnO (zinc oxide), 5-712, 10-173 および 2-1227 は α -Al₂O₃ (alpha aluminum oxide) である。また、6-662 は Os (osmium), 1-1244 は Zn (zinc) である。

次に成分数 $r = 3$ として、ZnO 5-664 を固定してみると表-4 のようになり、第3の成分の混合比は自然に小さくなって、主成分は二つであることがわかる。

なお、6-534 は NaPt₃O₄ (sodium platinum oxide), 3-195 は Mg Fe Si O₄ (silicate), 10-117 は Li₃ZrF₇ (lithium fluozirconate), 8-249 は Li₂TiO₃ (lithium titanate) である。

6. あとがき

ここに述べた方法は、結晶の同定分析を、電子計算機を用いて行なうことを目的としたものである。

同定分析を行なうためには、結晶学も X線解析学もそれほど深い知識を必要としないで、専門外の人々にとっても、簡単に利用できるべきではない。

情報検索に関していうならば、電子計算機は大量のデータを一般的に、迅速に処理するには適しているが、細部については、専門の人々の分析に及ぶものではない。

なお、今回の実験では、情報検索の一般論でよく用いられる呼出率 (recall factor) や適合率 (pertinency factor) 等の算出は、手許にあるデータがさほど多量ではなかったため行なわ

なかった。

また、この報告に述べた方法において、一次検索さえ有効に行なわれれば、二次検索については、山を長方形でなく、三角形や正規曲線で近似するなどいろいろな方法が考えられ、そうすればピークの位置だけを読み取って、幅の方は自動的にきめることもできる。

今回の実験は、極度に複雑な鉱物や、粘土のようにピークがめいりょうでなく円丘状をなすもの等を除いたものについて行なった。

したがって今後の方向としては、実用としては対象とする物質をたとえば鉱物関係に限って、それに応じた検索方法を開発することが必要であろう。

終わりに臨み、この研究を行なうに当たり、種々ご協力を賜わった地質調査所鉱床部鉱石課の諸氏および住鉱コンサルタント技師長 (元地質調査所鉱石課長) 大津秀夫氏に深く感謝いたします。

また、当研究の指針においてご指導をいただいた国立防災科学技術センター第3研究部長菅原正巳氏、結晶、X線関係においてお教えいただいた第2研究部地震防災研究室長高橋博氏、第2研究部大八木規夫氏に謝意を表します。なお、この研究を進めるに当たり、第3研究部尾崎書子、渡辺八重子両氏の協力をあおみだのこに感謝いたします。

(1968年1月8日 原稿受理,
1968年5月20日 改稿受理)